

DOCUMENT RESUME

ED 221 826

CS 006 809

AUTHOR Seminoff, Nancy Wiseman
TITLE Characteristics of Written Questions in Selected American History Textbooks: An Investigation.
PUB DATE Nov 81
NOTE 50p.; Paper presented at the Meeting of the National Council for the Social Studies (Detroit, MI, November 21-24, 1981).
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Content Analysis; *Inquiry; Junior High Schools; *Questioning Techniques; Reading Comprehension; Social Studies; Textbook Content; *Textbook Research; *United States History

ABSTRACT

Given the important role written questions have as a teaching strategy and their apparent use in textbooks, this study examined the written questions posed in representative American history textbooks intended for students at the seventh and eighth grade levels. Ten textbooks were randomly selected from a larger group of texts. Three chapters in each of the 10 texts were then randomly selected and the questions in each chapter identified. A seven-point scale was devised to describe the clarity, objectivity, and structure of each question. The same scale was further used to describe the extent to which each of the 30 sets of questions possessed sequence. In addition, each question was categorized by level of response--literal, interpretive, or applied. A summary of the means for the seven-point scale revealed the following scores: clarity--4.09, objectivity--6.18, structure--2.82, and sequence--4.09. Findings also revealed that literal questions were posed with considerably higher frequency in these selected books than interpretive and applied level questions. With the exception of a single text, applied level questions were excluded almost totally. The apparent absence of applied level questions and the extensive emphasis on literal level questions suggest that the questions are at odds with the emphasis of social studies on modes of inquiry. (HOD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED221826

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.

Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"Characteristics of Written Questions
in Selected American History Textbooks:
An Investigation"

National Council for the Social Studies

College and University Faculty Assembly
Research Roundtable Number Three
November 20, 1981

Nancy Wiseman Seminoff
Northern Michigan University
Marquette, Michigan

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Nancy Wiseman

Seminoff

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

CS006P09

"Characteristics of Written Questions in Selected
American History Textbooks: An Investigation"

Introduction

Formal schooling cannot provide students with sufficient knowledge to last a lifetime, but it can and must assist students in developing a means for continuing to acquire knowledge. The knowledge explosion requires continual reliance upon and refinement of the communication processes, reading, writing, speaking, listening. While reading is not a prerequisite to learning, reading does provide students with not only an important means by which to learn in school, but also with a medium by which they can continue their acquisition and application of knowledge.

Reading has long been considered the most important subject in the elementary school. And although reading has not traditionally been seen as an important subject for all secondary students, there has been an emphasis historically on the improvement of specific skills not previously mastered by "poor readers." Also, for proficient readers there has been an emphasis on increasing speed, expanding vocabulary, and refining study skills.

In recent years, however, educators have come to realize that instruction in the secondary schools should develop a student's ability to be selective and independent in acquiring new knowledge through print, to reject unsound ideas, and to draw reasonable con-

clusions. This realization has led to a broader interpretation of reading instruction in secondary education. Instruction in reading, it seems, cannot be solely the responsibility of elementary teachers or of a few highly-trained reading teachers in the secondary schools. Rather, reading instruction must be the responsibility of all teachers who use print as a means of conveying information (Herber, 1978; Beyer, 1977a).

The use of print material is widespread in secondary schools. The Report of the 1977 National Survey of Science, Mathematics, and Social Studies Education (Weiss, 1978) indicates that social studies teachers employ textbooks in 88 percent of the 7th-9th grade classes and in 89 percent of the 10th-12th grade classes. The report indicates further that "inadequate student reading ability causes a serious problem in grades 7-12 (p. 160)." Approximately one-half of the social studies teachers reported that inadequate student reading ability is a serious problem in their school. The responses of 44 of 298 principals for grades 7-9 and 50 of 270 principals for grades 10-12 show that in their schools, inadequate student reading ability is a serious problem in social studies.

According to authorities in social studies (Gross, 1977; Lunstrum, 1977; Lumstrum and Taylor, 1978) curricular reforms in the past decade have contributed to a new social studies with a more sophisticated conceptual structure and an emphasis on modes of inquiry. The infusion of new knowledge from anthropology, sociology, and other behavioral disciplines has increased the technical

vocabulary. The result of these influences has been that more is required and expected of students in reading and comprehending materials.

Beyer (1977a) refers to John Dewey's observation that the social studies are "the largest, most inclusive and most complex of all phenomena with which the mind has to deal" (p. 19). Consequently, Beyer submits, the skills of literacy are absolutely essential to the understanding of these phenomena.

Investigations by Braam and Roehm (1964) and Braam and Walker (1972) have revealed that secondary content teachers (including social studies teachers) were unaware of the reading skills required by their students to read effectively in the various disciplines.

Janz and Smith (1972) compared the readability levels of social studies materials with the reading ability of a sample of 590 students in grades eight to ten and found that: 1) all social studies texts ranged from readability levels of grade nine through college (Flesh formula), 2) for the eighth grade the texts were too difficult for more than 60% of the students, 3) for the ninth grade the texts were too difficult for more than 75% of the students, and 4) for the tenth grade the texts were too difficult for from 36% to 93% of the students.

The extensive use of print and the extensive investigation of the readability of that print in social studies have been established in the literature. The necessity to combine the content of social studies with the process to understand the content has been emphasized by social studies authorities (Beyer, 1977b; Fraenkel, 1973;

Herman, 1969).

Content teachers have a responsibility to assist students in understanding the print used to convey the concepts of their particular discipline. The research studies conducted at Syracuse University (Vacca and Herber, 1977) indicate classroom teachers with little formal knowledge of reading instruction are indeed "able to adapt reading/learning techniques to existing instructional routines without changing their perceived roles as teachers of content. Empirical evidence, combined with teacher and experimental observation, generally supports the notion that content teachers can effectively and efficiently teach reading through content" (p. 4).

This study concerns itself with an examination of questions which accompany textbook material. Specifically, this study examines the written questions posed in representative American history textbooks intended for students at the seventh and eighth grade levels.

Questions are an integral part of most instruction, whether they lead to structured recitation or open-ended inquiry. Research suggests that "as much as one-third of all classroom discourse consists of questions" (Hyman, 1970, p. 217). Questions are used as a key instructional tool for various reasons; students are familiar with the question-asking process; questions arouse interest and stimulate thinking; questions define the parameters of a field of study; and questions are used to direct discussion, clarify misunderstandings and check students' comprehension (Wease, 1976). Teachers' questioning behavior has been studied (Guszak, 1967;

Hoetker, 1968) and commented on, as have the types/classifications of questions (Bloom, 1954; Cornbleth, 1977; Fraenkel, 1973; Sanders, N., 1966/1971).

When print is employed as a medium of instruction, questions are often incorporated as a teaching strategy. The use of questions, both written and oral, is frequently associated with providing guidance in comprehension. The importance of questions related to print is evident from their presence in books, suggested use in teachers' manuals, and/or their observed use in the classroom. Herber (1970) asserts that combining follow-up analysis and discussion with questions posed in study guides promotes both reading comprehension and knowledge of content. Earle (1969) and Sanders (1970) make the same point.

Given the limitations of four years preservice teacher preparation, and the fact that many secondary teachers, especially at the junior or middle school level are now teaching outside of their major area of preparation, it seems only reasonable that teachers often rely heavily on the textbook material and the accompanying questions. One reason for employing the questions, it is assumed, is to provide assistance for the students in comprehending the assigned print. Given the important role written questions have as a teaching strategy and the apparent use of questions posed in textbooks, it seems imperative to examine the questions intended for use with textbook material. The study of text-provided questions for seventh and eighth grade students is especially necessary, it seems, because the students are expected to make the transition

from the highly structured instructional reading materials of the elementary grades to the subject-related materials in grades seven and eight and in the senior high school.

Clarity, objectivity, structure, and sequence are four of the characteristics of good questions identified by Sanders (1974a). Various levels of response--literal, interpretive, applied--have been identified also as assisting students in acquiring meaning from print (Herber, 1978). It is these specific characteristics and levels of response called for by questions that are employed in this study (Appendix A).

The purpose of this study was to determine the extent to which questions in representative social studies textbooks intended for students at the seventh and eighth grade levels possess the characteristics of effective questions as established in the professional literature, and call for responses which are literal, interpretive, or applied.

Methodology

Textbooks for the study were selected from the titles which appeared in El-Hi Textbooks in Print-1978 (1978). The ten texts chosen for the study were randomly selected from a larger group of texts, all of which had been previously determined to meet specific criteria (see Explanatory Notes). Three chapters in each of the ten texts were randomly selected, and the questions in each chapter were identified.

A seven-point scale, ranging from 1 (minimum presence of a characteristic) to 7 (maximum presence of a characteristic), was devised to describe each characteristic. Each question was judged

for each characteristic - clarity, objectivity, and structure. The same seven-point scale was used to describe the extent to which each of the 30 sets of questions possessed sequence. In addition, each question was categorized by level of response - literal, interpretive, or applied.

Three educators were trained over a four-month period to employ the scales for judging the characteristics of questions and for categorizing questions regarding a level of response. The data collection was conducted in three consecutive two-week periods immediately following this training period. The questions were assigned by chapter through a balanced randomized design to each of the judges.

The reliability for the three judges' responses regarding each characteristic, clarity, objectivity, structure, for the 1820 questions in the study was computed by using a two-way analysis of variance (Winer, 1971). The results indicate correlation coefficients as follows: clarity, .76; objectivity, .48; structure, .84; sequence, .65.

The grand mean for each characteristic was computed by combining the ratings of three judges for 1820 questions, each question contributing equal weight in the calculation. The text mean, on the other hand, was calculated by combining the judges' ratings for only the questions in that text.

The grand mean for clarity, objectivity, structure and sequence, and the mean for each of the ten texts for each characteristic was compared to an arbitrary established standard average rating of 4,

the mid-point of the seven-point scale. A paired-sample t-test was used to determine how many of the grand means and each text's means, on each of the four characteristics, were significantly different from the standard rating of 4. Then the binomial theorem was used to compute the probability that this many means would be significantly different by chance; this was done as a safeguard against the dangers of multiple significance testing. These tests were performed to give some initial and admittedly tentative indication about whether the questions in each text were high or low in each of the four characteristics.

In addition, a two-way analysis of variance was computed using texts and characteristics as independent variables. This test was performed to determine whether there were significant differences between texts, or characteristics, or an interaction of texts and characteristics. Where appropriate, this analysis of variance was followed up with a Newman-Keuls Contrast Test.

Findings and Conclusions

Clarity

The grand mean for clarity over three judges was 4.09 (Table 1). The results of the t test for clarity (Table 2) revealed that the grand mean for clarity was significantly (.05) above the standard average rating of four. This finding suggests that the questions in selected American history textbooks do possess clarity. However, examination of the t test results for the mean of each text (see Appendix C, Table A) revealed that the means of six texts (text one, five, seven, eight, nine and ten) were significantly below the stan-

Table 1
Summary of Means and Standard Deviations Over
Three Judges on Four Characteristics
of Questions

	Characteristics			
	Clarity	Objectivity	Structure	Sequence
Mean	4.09	6.18	2.82	4.09
SD	1.55	.81	1.40	1.55
N	5460	5455	5455	90

Table 2
Summary of t Tests for Three Judges' Combined
Ratings on Clarity, Objectivity, Structure
and Sequence for Ten Textbooks

Characteristic	t	df	Probability ^a
Clarity	5.3327	5459	.0001*
Objectivity	198.2552	5454	.0001*
Structure	62.1620	5454	.0001*
Sequence	.428	89	.3470 NS

^a computed two-tailed probability

*p. <.05

dard average rating while the means of only two texts, three and four, were significantly above.

Inspection of the data indicated the number of questions per text for these ten texts varied widely, ranging from a high of 535 questions in book three to a low of 67 questions in book six (Appendix B). The two texts whose means were above the standard average rating (Figure 1) contributed 702 questions to the grand mean ($N=1820$) while the six texts below the standard average rating contributed a total of 673 questions, establishing a clear possibility that the disproportionately high number of questions in just two texts may have adversely affected the grand mean.

Because of the wide variation in the number of questions per text and the fact that only two text means were above the standard average, a one-way analysis of variance was computed to determine if the means of the texts (calculated from the judges' combined ratings of clarity for each text) in fact differed from each other for clarity. Results of the analysis of variance tests of differences between the means of books for clarity revealed that the means did differ significantly ($.05$) (see Appendix C, Table E).

An analysis of variance by itself, however, can determine only that some texts are significantly different from others; it does not pinpoint which texts are significantly different. Therefore, the Newman-Keuls contrast test was employed to determine which text means differed significantly from each other (Appendix C, Table F).

		Book									
		1	2	3	4	5	6	7	8	9	10
7											
6					<u>5.46*</u>						
5				<u>4.63*</u>							
4											
3	<u>3.21*</u>	<u>3.98</u>			<u>3.70*</u>	<u>3.96</u>	<u>3.75*</u>		<u>3.90*</u>	<u>3.55*</u>	
2									<u>2.96*</u>		
1											
SD		1.01	.93	1.52	.88	.73	.82	.84	1.04	.85	.81

Figure 1: Profile of Means and Standard Deviations
Over Three Judges by Book for Clarity

*Significant at .05 level.

Results of the Newman-Keuls contrast test revealed that the means for clarity of books four and three were significantly higher than the means of the other eight books, and also that book four's mean was significantly higher than book three's.

The significant difference in the means for clarity for books three and four, and the extensive number of questions in the two books combined, seems to indicate that these two books were contributing disproportionately to the grand mean for clarity that was significantly above the standard rating, suggesting that books three and four may have adversely influenced the grand mean for clarity.

It is the investigator's conclusion that the finding which indicated the grand mean of the texts was significantly above the standard average of four is probably a function of two texts contributing to that significance. The fact that the means of six texts were significantly below the standard average rating indicates that questions in these texts are not characterized by clarity to a very great extent.

Three related conclusions can be drawn. 1) Although a generalization regarding all American history texts must be viewed cautiously, the random selection of texts for this study argues that other texts randomly selected would follow the same pattern of more than half of the texts containing questions which lack clarity. 2) Random observation of the types of questions posed in texts three and four suggest that the questions judged to possess clarity frequently directed the student to choose from multiple responses, fill in the blank with a single word or phrase, or match a term or name with a definition or phrase. Questions which appeared to lack

clarity were types which inquired if the student knew the meaning of specific words and/or which required the student to respond to more than one task. 3) As reflected in the acceptable inter-rater reliability (.76), clarity is a construct which can be employed successfully by a group of judges trained in its use.

Objectivity

The grand mean for objectivity over three judges was 6.13 (Table 1). The results of the t test for objectivity (Table 2) revealed that the grand mean for objectivity was significantly (.05) above the standard average rating of four. Examination of the results of a t tests for the mean of each text (Appendix C, Table B) revealed that all ten of the texts were significantly above the standard average rating. These findings suggest that questions in these texts are characterized by objectivity.

The investigator's concern for the cause of the low inter-rater reliability led to further examination of the data for objectivity. This examination revealed the frequent choice by the judges of points 5, 6, and 7 on the seven-point scale. This extensive selection of points 5, 6, and 7 in effect reduced the seven-point scale to a three-point scale, i.e., "shrinking" the scale. Mathematically, the fact that the means of the ratings of the individual judges were so close to the maximum value (i.e., 7) caused small differences between the means resulting from the judges' ratings to assume more importance than they might have elsewhere. In addition, one effect of "shrinking" the scale is that ratings which in other situations would be normal now become "outliers" and, therefore, distort the

reliability coefficients.

The investigator thus concluded: 1) that questions in selected American history textbooks are characterized by objectivity; this finding was at least in part the function of the large number of literal level questions posed in the texts selected for this study; 2) that the judges appear to have agreed in the use of the definition of objectivity to a greater extent that is indicated by the inter-rater reliability; 3) that although a generalization regarding all American history textbooks must be viewed cautiously, the random selection of texts for this study argues that other texts randomly selected would follow the same pattern of containing questions which are characterized by objectivity according to the definition posed in this study.

Structure

The grand mean for structure over three judges was 2.82 (Table 1). The results of the t test for structure (Table 2) revealed that the grand mean for structure was significantly below the standard average rating of four. Examination of the t tests for the mean of each text (Appendix C, Table C) revealed that nine of the tests were significantly below the standard average rating and one (book four) was significantly above. These findings indicate that questions in the majority of these texts do lack structure.

The noticeably high mean of the judges' ratings for the questions in text four caused the investigator to examine the data for this text more closely. Closer inspection of the questions in this text indicated that they were frequently of the type that required

the student to fill in the blank, choose from multiple responses, and/or match a term or person's name with a definition or phrase. Examination of the judges' ratings for the questions in book four further indicated that the types of questions the judges considered to be structured (by ratings of 5, 6, or 7) were frequently the types indicated here.

Several conclusions can be drawn. 1) Questions which require the student to fill in the blank, choose from multiple responses, and to match terms and/or phrases clearly seemed in the opinion of the judges to offer more assistance or guidance (more structure) than did questions which merely direct students to "discuss" or "write" about individual terms. 2) The high inter-rater reliability (.84) for structure indicated that the judges perceived the definition similarly in employing it in this study. It seems structure is a construct which can be employed successfully by a group of judges trained to use it in judging questions. 3) Questions in the selected American history textbooks examined in this study did lack structure. Although a generalization regarding American history textbooks intended for seventh and eighth grade must be viewed cautiously, the random selection of texts for this study argues that the questions in similar texts might likewise be assumed to lack structure.

Sequence

The grand mean for sequence across three judges was 4.09 (Table 1). The results of the t test for question five revealed that the grand mean for sequence (Table 2) was neither significantly above nor significantly below the standard average rating of four.

Examination of the t test results for the means of each text (Appendix C, Table D) revealed that the means of three texts (texts two, six and ten) were significantly above the standard average rating and the mean of one text (text seven) was significantly below.

The limited number of significant means called for a closer inspection of the questions in texts two, six, seven, and ten (the texts with means either significantly higher or significantly lower than the "standard" of four). Sequence is defined as the relationship of questions and/or parts within a question to each other in such a manner that the response to each in the order presented assists in developing increasingly complex understandings beyond the literal level. This definition suggests a relationship between the levels of response and sequence.

Inspection of the data (Table 2) indicated that books two, six and seven contain a high proportion of interpretive level questions (65.3%, 72.7%, and 71.6%, respectively). In contrast, the other books contain proportions which range between 18% and 45.6%. Book ten, although containing a proportion of only 44.1% interpretive questions, had the highest proportion of applied level questions (43.2%).

A common factor in the six texts containing means which were non-significant was their high frequency of literal questions. It appears from the data that one is unlikely to find sequence, as defined in this study, in a set of mostly literal questions. The mere presence of interpretive and/or applied level questions, on

Table 3
Frequency Distribution for "Level-of-Questions"
Categories According to Judges Majority
Classification by Book and by Category

Categories	1	2	3	4	5	6	7	8	9	10	Total
Literal											
Number of Resp.	49	12	280	98	124	8	11	118	129	12	841
Adjusted Freq. %	(51.6)	(5.9)	(53.2)	(59.0)	(68.1)	(12.1)	(11.6)	(71.1)	(77.3)	(10.8)	(47.3)
Interpretive											
Number of Resp.	38	134	240	46	54	48	68	40	30	49	747
Adjusted Freq. %	(40.0)	(65.3)	(45.6)	(27.7)	(29.7)	(72.7)	(71.6)	(24.1)	(18.0)	(44.1)	(42.0)
Applied											
Number of Resp.	7	49	6	10	3	10	16	8	7	48	164
Adjusted Freq. %	(7.4)	(23.9)	(1.2)	(6.0)	(1.6)	(15.1)	(16.8)	(4.8)	(4.2)	(43.2)	(9.2)
Unclassifiable											
Number of Resp.	1	10	0	12	1	0	0	0	1	2	27
Adjusted Freq. %	(1.0)	(4.9)	(0)	(7.2)	(.5)	(0)	(0)	(0)	(0)	(1.8)	(1.5)
No Agreement ^a	1	7	9	1	1	1	0	0	2	19	41
Totals	96	212	535	167	183	67	95	166	169	130	1820

^aThe "No Agreement" cases were eliminated from the computation of the adjusted frequencies.

the other hand, does not insure sequence, as is clearly demonstrated by the judges' ratings of the questions in book seven.

Though beyond the scope of this study, it is interesting to contemplate the relationship of levels and sequence. One compelling question is whether the literal questions in texts two, six, and ten differ significantly from the literal questions in text seven in structure (a structured question conceivably contributing more to the development of "increasingly complex understanding" than an unstructured one); if they differ only in their relationship to the total set; or if the difference in sequence is more a reflection of the relationships among interpretive and applied level questions than of the relationship of literal level questions to either or both of these.

It is of course possible, given the inter-rater reliability of .65, that there was a lingering confusion in the minds of the judges regarding the intent of the definition and/or its use in rating sets of questions. It should be further noted that the complexity of assessing the inter-relatedness of questions within the "whole," demanding a Gestalt view of individual chapters, might itself have posed difficulty. In addition, the N of 30 for sequence, a substantially lower N than that for the other characteristics (clarity, objectivity, structure), might have contributed to the lack of statistical significance.

Literal, Interpretive and Applied Levels

47.3% of the 1779 questions agreed upon by at least three judges called for responses which were literal, 42% of the same

questions called for responses which were interpretive, and 9.2% called for responses which were applied (Table 3). The number and proportion in each category were computed by combining the questions in the ten texts. Closer inspection of the data, however, reveals that the questions in one or two texts unduly influenced the total number of questions in each category. Book three (which contained 535 questions) contributed 280 questions to the total of 841 literal questions, i.e., 33% of the literal questions identified in the study. Book three also contributed 240 questions to the total of 747 questions classified as interpretive, i.e., 32% of the interpretive questions in the study. Books ten and two together contributed 97 questions to the total of 164 applied level questions, i.e., 59% of the applied level questions in the study. Therefore, it appears that in each category one or two books contributed a disproportionately large number of the total sample of questions.

Further inspection of the data for each text by level reveals that the proportion of questions by level within a text is not reflected by the total proportion of questions across texts for that level. Although the total proportion of the questions across texts at the literal level was 47.3%, six of the texts (books one, three, four, five, eight, and nine) contain a proportion of literal level questions in excess of 50% (51.6%, 53.2%, 59%, 68.1%, 77.3%, respectively). The remaining four texts contain a considerably lower proportion of literal questions, that proportion ranging from 5.9% to 12.1%. Clearly, there is a preponderance of literal level

questions in most of the texts examined in this study, a finding not apparent in the proportion of literal questions across texts.

Of the four texts containing proportionately few literal level questions, three (books two, six, and seven) contain a proportion of interpretive level questions in excess of 50%, that proportion ranging from 65.3% to 72.7%. Further inspection reveals that the combined total of the interpretive level questions in these three texts is 33% of the 747 interpretive level questions identified in the study. Clearly there is a limited proportion of interpretive level questions in individual texts, a conclusion not apparent from the consideration of the total proportion of interpretive questions reported.

Text ten contains an almost equal proportion of interpretive and applied level questions (44.1% and 43.2%, respectively). This text also contains the highest number and proportion (19 questions or 46%) of the questions not agreed upon by at least two of the judges. Casual inspection of the questions in text ten reveals that they are more complex (i.e., lengthier and/or more conceptually demanding) than the questions in the other nine texts. The fact that the judges agreed on the classification of numerous applied level questions in this text suggests that it was the complexity of the questions rather than the definition of applied level questions that was the cause of the judges' difficulty.

It is the investigator's conclusion that insight gained from the total number or proportion of questions in a category across texts is not as valuable as insight gained from examination of the number or proportion of questions within individual texts. One or

two books, it is clear from the present study, may unduly influence the proportion of questions across texts within a level. A second conclusion is that literal level questions are posed with considerably higher frequency in selected American history books than interpretive and applied level questions, applied level questions, with the exception of a single text, being excluded almost totally.

Implications

Specific characteristics of effective written questions and levels of response called for by questions were identified for this study. The results of the investigation clearly indicate that these characteristics and levels can be successfully employed as constructs to describe questions posed in selected American history textbooks intended for seventh and eighth grade students. Presumably, these same constructs can be used to describe the questions posed in texts employed in other disciplines and/or for other grade levels.

There is evidence in the investigation to suggest that these constructs are not independent--that there may be an interrelatedness among the characteristics and/or levels of questions. For example, the questions in two books (books three and four), which were judged to possess clarity to a great extent, were also judged to possess objectivity to a great extent. Similarly, in several texts, questions which were categorized as literal tended to be judged objective. The extensive proportion of interpretive and/or applied level questions in three texts (books two, six, and ten), and the apparent sequence within sets of questions in those texts, suggests a possible relatedness of these two constructs as well. The fact

and the extent of interrelatedness of characteristics and/or levels is not, of course, made clear by the present study.

Intriguing evidence in the present study is the apparent inter-relatedness between sequence and level, a point commented upon in the preceding section. The present study considered four of the characteristics of good questions presented by Sanders (1974a)--clarity, objectivity, structure, and sequence--but by design did not include the characteristic of focus. Sanders states that good questions should "focus on selected aspects or dimensions of meaning which when considered lend a sense of the whole" (p. 80). This suggests that good questions originate with a conscious determination by the questioner of the particular purposes the questions are to serve, especially their relationship to a specific content.

It does seem possible, at least theoretically, for a given set of questions to possess all of the characteristics examined to some or a considerable degree but still not be worthwhile. Henry's (1974) comments on the importance of a synthesis of understandings/concepts from the consideration of several sources or selections argue that it is through the planned consideration of these several selections that broader meaning and insights extending beyond the literal level and beyond any one of the selections evolve. Henry's notion of synthesis seems clearly analogous to the notion of levels employed in the present study. The apparent absence of applied level questions and the extensive emphasis on literal level questions suggest that questions in these texts are curiously at odds with the purpose of social studies as defined in the professional literature. It is tempting to conclude on the basis of the texts ex-

amined that there is an apparent confusion in distinguishing between the texts and the curriculum. The questions in the texts examined for this study seem only infrequently to involve the students in the process, particularly the decision-making process, necessary to establish and support particular understanding(s) and/or point(s) of view.

The present study does not offer insight regarding which of the four characteristics considered is the most important, nor does it suggest that those characteristics employed constitute the universe of characteristics of effective questions. Nevertheless, because educators do purchase instructional materials which include questions as an instructional aid, those questions must be examined carefully. It is the expectation of this investigator that the relative absence in the texts examined of the characteristics identified in the study, and the apparent absence of focus, would most likely be true of the texts employed in other disciplines and/or at other grade levels.

Suggestions for Further Research

This study has attempted to define a theoretically-based model for describing questions posed in selected American history textbooks intended for seventh and eighth grade students. The findings of the study seem clearly to indicate that the characteristics and levels can be employed successfully to describe questions in these texts, and presumably the questions provided for use with other disciplines and/or other grade levels.

The study thus provides a reference point for other investigations, in particular other investigations which explore the quality

of questions and the intended purpose of those questions as instructional aids. Several promising directions for further research are listed below:

1. It would be desirable to replicate this study with questions in texts employed in other disciplines and/or at other grade levels.
2. The possible interrelatedness among characteristics and/or levels should be investigated. The need to investigate the potential relationship between levels and sequence, and particularly the relationship between clarity, objectivity, and structure, seem particularly compelling.
3. Examination of sequence within several sets of questions within a text should be conducted to allow for a fuller perception of the extent of sequence in a text than was possible in this study.
4. Questions, either those provided in texts or those generated by the teacher, should be examined by considering the characteristics and levels presented in this study in conjunction with the intended purpose, i.e., the focus, of those questions.
5. Assuming the clarification of the intended purpose of focus of questions and of the relationship between levels and the four characteristics, methodological studies should be conducted to determine the effect of questions on students' understanding of material they read.

Explanatory Notes

Criteria for Selecting Texts:

1. the text had to be indicated as being designed for use in grades seven and/or eight;
2. a teacher's annotated edition or a teacher's guide had to accompany the student edition available for purchase;
3. the text had to be available in a single volume or the hardcover;
4. there had to be an indication of comprehensive content coverage, i.e., the text content had to cover a broad time span or several topics;
5. the text had to have been published sometime between 1974 and 1979;
6. there had to be provision of questions within the student's edition of the text.

APPENDIX A

DEFINITIONS

Question - an item which appears in writing and is intended to elicit a response with regard to the material it accompanies. Questions included for examination in this study are those items which call upon the reader to relate and/or utilize the information, concepts and/or generalizations drawn from the material accompanied by the item(s) in completing the stated task. Such items may be located in the initial portion of the print, at various points throughout the print, and/or at the conclusion of the print presentation.

Items excluded from examination are those which 1) are distinctly and physically set apart from the chapter presentation and identified as: a) a test, b) a skills lesson, c) a bibliography; 2) accompany graphic aids where explicit reference to these aids is not made in the print presentation.

For the purpose of this study levels of questioning are defined as follows:

Literal - requires the respondent to focus on evidence provided in the selection; does not require the discernment of meanings not stated explicitly in the print. Examples of literal questions:

1. According to the text, what qualities did Truman have as President?
2. List four personal qualities which, according to the text, Truman displayed as President.

Interpretive - requires the respondent to discern meanings not stated explicitly in the selection, through the joining together of literal evidence to form generalizations or inferences, and/or to sense implications based on inferences previously discerned from the selection, rather than reporting evidence explicitly stated in the selection. Examples of interpretive questions:

1. How in your view did Truman's personal characteristics influence his performance as President?
2. Do you think, according to the evidence in the text, that the roles of president and the two houses in the lawmaking process illustrate the principle of checks and balances? Explain.

Application - requires the respondent to discern meaning through a synthesis of evidence from within the selection and external to it; affords clear potential for the evolution of new ideas/relationships that encompass both types of evidence but extend beyond them. Judgment regarding this evidence is thus "selection free" because the evidence employed in making such judgment is both from within and outside the selection. Examples of application question:

1. Eisenhower did not approve of the Brown decision, but enforced it to the best of his ability. Compare his action with that of Andrew Jackson on the Cherokee Claims (pages 191-192). Do you feel that either President acted correctly? Why? Defend your answer.

For the purpose of this study characteristics of effective questions are defined as follows:

Clarity - The question is specific in indicating the task(s) to be completed by the respondent. The language employed in the question is precise rather than vague. Examples of questions which possess clarity:

1. List specific advantages and disadvantages of having all workers with the same skill in one union.
2. Name the new national labor union that was organized in 1935.
3. Decide which of the following you would expect to find in Montana. Circle your choice(s). Pages 15-20 may help you.

sheep
grazing

cattle
corn

sugar beets
farming

Examples of questions which lack clarity:

1. Do you see any advantage in having all workers with the same skill in one union? Do you see any disadvantages?
2. Identify these dates: 1776, 1812, 1939.
3. Discuss the meanings of these words:
romantic opportunities
4. Would you say that the balance of power in Europe (and India and the New World) had been upset by Britain's victory? What happens when one nation is too powerful?

Objectivity - The phrasing of the question allows the respondent to provide his/her own response rather than being restricted to supporting the stated bias of the author of the question. While the respondent may be required to generate or identify evidence to support the response, his/her own perception is encouraged. Examples of questions which possess objectivity:

1. What effect, if any, do you think draft laws had on the growth of families during the war? After the war?
2. Write a statement which reflectes your thinking on the issue of government control over TV programming.

Examples of questions which lack objectivity:

1. Why should government controls over television programming be made more strict?
2. What is the main idea of Chapter 19?
 - A. to describe the beginning of a new nation
 - B. to review the causes of the American Revolution
 - C. to describe the meeting of the Second Continental Congress

Structure - The guidance or help 'built into' the question that assists the respondent in completing the stated task. This guidance might assist the respondent in establishing essential information; in drawing inferences; or in applying sophisticated statements of understanding to related understandings beyond the selection, i.e., synthesizing understandings. This guidance might consist of location aids: page number, column, paragraph, line, etc.; or words, phrases, sentences to be manipulated or reacted to. The various forms of guidance can be provided at any level of questioning. Examples of questions which possess structure:

1. Complete both parts A and B in deciding on the advantages and disadvantages of big business to the consumer.
 - A. Identify the advantages and disadvantages of big business to the consumer by placing each word/phrase in the appropriate column:

decrease quality
mass production
competitive prices
fixed prices

ADVANTAGES

monopoly
variety of choice
increased availability

DISADVANTAGES

B. On the basis of your responses in Part A, select the sentence(s) below which you think are reasonable conclusions. Use an X to indicate your choice(s).

- ☐ 1. Mass production can allow the consumer a greater variety of choice of items.
- ☐ 2. Mass production can result in decreased quality of items.
- ☐ 3. A monopoly can result in the elimination of competitive prices.
- ☐ 4. Other _____

Examples of questions which lack structure:

- 1. What is democracy?
- 2. Who are the following: John F. Kennedy, Abraham Lincoln?
- 3. On what policies could Americans agree?

Sequence - In this study sequence is defined as the relationship of questions and/or of parts within a question to each other in such a manner that the response to each in the order presented assists the respondent in developing increasingly complex understanding beyond the literal level of information provided in the text.

Examination of numerous social studies textbooks by this writer reveals that questions in such textbooks generally appear in one or more of three locations: preceding the print presentation; contained within the print presentation; or following the print presentation. Although there may be a discernible sequence within the subsets of questions incorporated in the identified chapters, to consider the matter of sequence in such subsets is to risk tampering with the author's intention. For purposes of this study, sequence is to be considered only with reference to all questions in the identified chapter. Examples of questions which possess sequence:

- 1. List three ways, according to the text, that the Plains Indians appeared to depend on the buffalo.
- 2. List two reasons that the white man hunted buffalo.
- 3. State how the Plains' Indians' way of life seemed to be affected by the white man hunting the buffalo.

4. Discuss how the Indians reacted to the white man hunting the buffalo.
5. Do you feel the actions of the Indians were justified or not justified? State a reason for your decision.

Examples of questions which do not possess sequence.

1. Do you feel the actions of the Indians were justified or not justified? State a reason for your decision.
2. Discuss how the Indians reacted to the white man hunting the buffalo.
3. List three ways, according to the text, that the Plains Indians appeared to depend on the buffalo.
4. State how the Plains Indians' way of life seemed to be affected by the white man hunting the buffalo.

APPENDIX B

INFORMATION REGARDING CHAPTERS AND QUESTIONS SELECTED FOR STUDY

Text Number and Title	Chapter Number	Actual No. of Chapter in Text	Pages per Chapter	Total Pages per Chapter	Total Pages per Text ^a	% of Text Pages Selected	Questions Designated in Study per Chapter	Total Questions Designated in Study per Text
1. America: America	1 2 3	6 33 45	15 9 9	33	683	5%	43 28 25	96
2. American Dream	4 5 6	7 8 13	37 26 38	101	695	15%	78 60 74	212
3. American History	7 8 9	8 10 12	39 49 51	139	697	20%	135 216 184	535
4. Exploring Our Nation's History	10 11 12	9 19 34	13 16 18	47	705	7%	51 44 72	167
5. Foundations of Freedom	13 14 15	3 16 20	12 21 18	51	516	10%	32 80 71	183
6. Impact of Our Past	16 17 18	3 9 17	25 25 37	62	799	8%	23 22 22	67
7. Pageant of American History	19 20 21	7 17 27	27 21 21	69	689	10%	35 31 29	95
8. This Is America's Story	22 23 24	1 14 27	22 24 14	60	729	8%	66 58 42	166
9. Two Centuries of American Progress	25 26 27	14 18 28	20 20 19	99	694	9%	61 68 40	169
10. United States Past and Present	28 29 30	3 9 13	32 32 27	91	448	20%	33 57 40	130

^aRefers to total pages of chapters but excludes reader aids, e.g., maps, glossary, at end of text.

APPENDIX C

Table A

Summary of t Tests for Combined Ratings of Three
Judges for Each of Ten Textbooks for Clarity

Textbook Number	t	df	Actual two-tailed Probability
1	13.262	287	< .0001*
2	.551	635	.8222 NS
3	16.552	1604	< .0001*
4	37.152	500	< .0001*
5	9.483	548	< .0001*
6	.687	200	.9996 NS
7	5.018	284	< .0001*
8	22.377	497	< .0001*
9	2.623	506	.0176*
10	<u>10.927</u>	<u>389</u>	< <u>.0001*</u>
Total	5.333	5459	< .0001*

*p < .05

Table B
Summary of t Tests for Combined Ratings of Three
Judges for Each of Ten Textbooks for Objectivity

Textbook Number	t	df	Actual two-tailed Probability
1	58.372	287	<.0001*
2	84.213	635	<.0001*
3	130.818	1599	<.0001*
4	68.843	500	<.0001*
5	57.511	548	<.0001*
6	25.485	200	<.0001*
7	26.171	284	<.0001*
8	55.353	497	<.0001*
9	88.936	506	<.0001*
10	37.349	389	<.0001*
Total	198.163	5454	.0001*

*p < .05

Table C
Summary of t Tests for Combined Ratings of Three
Judges for Each of Ten Textbooks for Structure

Textbook Number	t	df	Actual two-tailed Probability
1	45.191	287	$< .0001^*$
2	23.391	635	$< .0001^*$
3	13.224	1599	$< .0001^*$
4	9.628	500	$< .0001^*$
5	56.433	548	$< .0001^*$
6	37.529	200	$< .0001^*$
7	30.050	284	$< .0001^*$
8	76.808	497	$< .0001^*$
9	61.375	506	$< .0001^*$
10	<u>38.724</u>	<u>389</u>	$< \underline{.0001^*}$
Total	62.133	5454	$< .0001^*$

* $p < .05$

Table D
Summary of t Tests for Combined Ratings of Three
Judges for Each of Ten Textbooks for Sequence

Textbook Number	t	df	Actual two-tailed Probability
1	2.286	8	.9998 NS
2	5.547	8	.0008*
3	1.315	8	.2236 NS
4	1.069	8	.6344 NS
5	1.941	8	.1718 NS
6	3.773	8	.0112*
7	7.249	8	.0008*
8	.632	8	.90 NS
9	1.048	8	.6528 NS
10	8.222	8	.0002*
Total	.428	89	.3470 NS

*p < .05

Table E
Analysis of Variance Summary Table for Clarity
Between Ten Books

Source	SS	df	MS	F	Sig.
Between Books	2526.003	9	280.667	228.556	<.0001*
Within Books (Error)	6690.651	5450	1.228		
Total	9216.654	5459			

*p < .05

Table F
Differences Between Pairs of Means for Books on Clarity
With Significance Levels for Those Differences

Book		Four	Three	Two	Six	Nine	Seven	Five	Ten	One	Eight
	Mean	5.461	4.627	3.980	3.960	3.901	3.751	3.703	3.554	3.208	2.956
Four	5.461	-	0.834	1.481	1.501	1.560	1.710	1.758	1.907	2.253	2.505
Three	4.627	.01	-	0.647	0.667	0.726	0.876	0.924	1.073	1.419	1.671
Two	3.980	.01	.01	-	0.020	0.079	0.229	0.277	0.426	0.772	1.024
Six	3.960	.01	.01	ns	-	0.059	0.209	0.257	0.406	0.752	1.004
Nine	3.901	.01	.01	ns	ns	-	0.150	0.198	0.347	0.693	0.945
Seven [†]	3.751	.01	.01	ns	ns	ns	-	0.048	0.197	0.543	0.795
Five	3.703	.01	.01	ns [†]	ns	ns	ns	-	0.149	0.495	0.747
Ten	3.554	.01	.01	ns	.05	ns	ns	ns	-	0.346	0.598
One	3.208	.01	.01,	.01	.01	.01	.01	.01	.05	-	0.252
Eight	2.956	.01	.01	.01	.01	.01	.01	.01	.01	ns	-

Books are listed in rank order from highest to lowest.

Significance levels are based on Newman-Keuls contrast tests (Winer, 1971, 185-196; 210-220)

BIBLIOGRAPHY

- Beyer, B.K. Back-to-basics in social studies. Boulder, CO: Social Science Education Consortium, 1977 (a).
- Beyer, K. Teaching basics in social studies. Social Education, 1977, 41, 96-104 (b).
- Bloom, B.J. Taxonomy of educational objectives, handbook I: the cognitive domain. New York: McKay, 1954.
- Braam, L.S. and Roehm, M. Subject area teachers' familiarity and reading skills. Journal of Reading, 1964, 8, 188-196.
- Braam, L.S. and Walker, J.E. Subject awareness of reading skills. Journal of Reading, 1972, 16, 608-611.
- Cornbleth, C. Using questions in social studies. How to do it notebook, series 2, #4. Arlington, VA: National Council for the Social Studies, 1977.
- Earle, R.A. Developing and using study guides. In H.L. Heber and P.L. Sanders (Eds.), Research in reading in the content areas, first year report. Syracuse, NY: Reading and Language Arts Center, Syracuse University, 1969.
- El-Hi Textbooks in Print - 1978. New York: Bowker, 1978.
- Fraenkel, J.R. Helping students think and value. Englewood Cliffs, NJ: Prentice-Hall, 1973.
- Gross, R.E. The status of the social studies in the public schools of the United States: facts and impressions of a national survey. Social Education, 1977, 41, 194-200, 205.
- Guszk, F.J. Teacher questioning and reading. Reading Teacher, 1967, 21, 227-234.

- Henry, G.H. Teaching reading as concept development: emphasis on affective thinking. Newark, DE: International Reading Association, 1974.
- Herber, H.L. Teaching reading in content areas. Englewood Cliffs, NJ: Prentice-Hall, 1970.
- Herber, H.L. Teaching reading in content areas (2nd edition). Englewood Cliffs, NJ: Prentice-Hall, 1978.
- Herman, W.L. Reading and other language arts in social studies instruction: persistent problems. In R.C. Preston (Ed.), A new look at reading in the social studies. Newark, DE: International Reading Association, 1969.
- Hoether, J. Teacher questioning behavior in nine junior high school English classes. Research in the Teaching of English, 1968, 2, 99-108.
- Hyman, R.T. Ways of teaching. New York: Lippincott, 1970.
- Janz, M and Smith, E. The students' reading ability and the readability of secondary school subjects. Elementary English, 1972, 49, 622-624.
- Lunstrum, J.P. Reading in the social studies. National Council for the Social Studies Yearbook, 1977, 47, 109-139.
- Lunstrum, J.P. and Taylor, B.L. Teaching reading in the social studies. Urbana, IL: ERIC Clearinghouse on Reading and Communication-Skills, 1978.
- Sanders, P.L. An investigation of the effects of instruction in the interpretation of literature on the responses of adolescents to selected short stories. Unpublished doctoral dissertation, Syracuse University, 1970.

- Sanders, P.L. When teachers talk and students (sometimes) listen. English Journal, 1974, 63, 80-81.
- Sanders, N.M. Classroom questions, what kinds. New York: Harper, 1966.
- Sanders, N.M. New questioning strategies build into United States history textbooks. Social Education, 1971, 35, 770-775.
- Vacca, R.T. and Heber, H.L. Research in reading in the content areas: a perspective. In H.L. Heber and R.T. Vacca (Eds.), Research in reading in the content areas. Syracuse, NY: Reading and Language Arts Center, Syracuse University, 1977.
- Wease, H. Questioning the genesis of teaching and learning. High School Journal, 1976, 59, 258-265.
- Weiss, I. Report of the 1977 national survey of science, mathematics and social studies education (Vol. 2). Washington, DC: U.S. Government Printing Office, 1978.